Appendix 1

Deep learning (DL) model architecture, loss function, and data splitting

The DL model architecture is shown in *Figure S1*. More details about the architecture can be found in (18,30).
The DL model was trained in a supervised manner with the loss function:

$$L = \left|\left|T1_{DL} - T1_{MAGiC}\right|\right|_{l1} + \left|\left|T2_{DL} - T2_{MAGiC}\right|\right|_{l1} + \left|\left|PD_{DL} - PD_{MAGiC}\right|\right|_{l1}$$

As can be seen, the loss function is composed of the sum of three terms, one per output parametric map [i.e., longitudinal relaxation time (T1), transversal relaxation time (T2), and proton density (PD) maps]. Each term consists of the l1 norm, also known as mean absolute error (MAE), between synthesized and the reference parametric maps. Note that in this case the reference parametric maps are the magnetic resonance image compilation (MAGiC) maps. MAE is defined as the mean of the absolute difference between both maps.

Finally, *Table S2* shows the patient indexes included in training, early-stopping validation and test sets for each split.

Voxel-wise statistical prediction of T1w-enhancement

Nunez-Gonzalez *et al.* (13) proposed a method for voxel-wise statistical prediction of normal/ABN tissue and tissue with and without T1e from only pre-contrast MAGiC parametric maps. In that study, the authors performed the voxel-wise prediction using a receiver operating characteristic curve (ROC) analysis. The authors considered three classification problems: classification-I (C-I) abnormal tissue (ABN) *vs.* normal white matter (nWM); classification-II (C-II) T1e *vs.* (nWM + T2h); and classification-III (C-III) T1e *vs.* T2h (only inside ABN). They also considered four voxel-wise metrics: (I) T1 values, (II) T2 values, (III) normT1T2 (i.e., the Euclidean norm of the T1 and T2 values), and (IV) normlog (i.e., the Euclidian norm of the logarithm of T1 and T2 values). For each classification problem the metric with the highest area under the curve (AUC) was selected. The thresholds for optimal classifications were calculated as the highest Youden's index of the ROC curve. The selected metrics and thresholds were normlog with threshold 8.44 for C-I, normT1T2 with threshold 1,344 ms for C-II, and normT1T2 with threshold 1,512 ms for C-III. The authors showed that the ROC analysis including PD did not improve the AUC in any of the cases. Consequently, they excluded the PD values from the rest of the analysis. Accordingly, we perform the classifications by applying their selected thresholds to the aforementioned metrics inside the white matter plus ABN masks to discriminate between the different regions.

Additional qualitative analysis

Figure S2 shows a representative slice of both MAGIC and synthesized maps for all patients of GLIOMA dataset. It can be noticed that most of the differences between both maps are located at the interfaces between tissues. This effect might be caused by partial volume effects and misregistration of the input weighted images. In general, PD maps present less differences than T1 and T2 maps for all patients. In patients 2 and 6 the T1 values in the resected regions are higher in the synthesized than in the MAGiC maps. We hypothesize this could be due to a poor estimation of the MAGiC maps in these two patients, given that these extremely low T1 values in the resected regions are not present across the other 12 patients.

Figure S3 shows the segmented T1e and T2h regions obtained through the statistical prediction from both MAGiC and synthesized maps for all patients of GLIOMA dataset. It is worth noting that segmented T1e and T2h regions are similar in both maps for all patients and that both present an overestimation of the T1e with respect to HD-GLIO. We hypothesize this could happen due to an altered interstitial fluid mobility and increased water content in the perivascular space.

https://dx.doi.org/10.21037/qims-24-721

Table S1 Details about the acquisition parameters for the different image modalities and each dataset

	Pre-T1w	T2w	T2w-FLAIR	Post-T1w	MAGiC	T1 map	T2 map	PD map
GLIOMA								
Voxel size (mm^2)	1.0×1.0	0.6×0.6	1.1×1.1	1.0×1.0	1.0×1.0			
Slice thickness (mm)	1.0	3.0	1.6	1.0	3.0			
Slice spacing (mm)	–	3.0	–	–	3.0			
# Slices	352	49	224	352	49			
FOV (mm^2)	240×240	233×233	246×246	240×240	240×240			
TE (ms)	3.3	97	89	3.3	6,114			
TR (ms)	7.9	9,837	5,000	7.9	15.7			
TI (ms)	450	–	1,588	450	11			
Flip angle (°)	12	90	90	12	90			
Scan time (min)	~5	~4	~4	~5	~5			
UPenn-GBM								
Voxel size (mm^2)	0.98×0.98	0.9×0.9	0.94×0.94	0.98×0.98				
Slice thickness (mm)	1.0	0.9	3.0	1.0				
Slice spacing (mm)	NA	NA	NA	NA				
# Slices	155	155	155	155				
FOV (mm^2)	240×240	240×240	240×240	240×240				
TE (ms)	3.1	458	140	3.1				
TR (ms)	1,760	3,200	9,420	1,760				
TI (ms)	950	–	2,500	950				
Flip angle (°)	15	120	170	15				
RMaps								
Voxel size (mm^2)	1.25×1.25	1.02×1.36				1.5×1.5	1.5×1.5	1.5×1.5
Slice thickness (mm)	1.2	3				1.5	1.5	1.5
Slice spacing (mm)	–	3				–	–	–
# Slices	170	50				150	150	150
FOV (mm^2)	240×240	260×195				240×240	240×240	240×240
TE (ms)	3	85				2	17, 46, 75, 104, 133, 162	2
TR (ms)	6.44	4,000				18	1,000	50
TI (ms)	900	–				–	–	–
Flip angle (°)	10	90				2, 3, 4, 5, 7, 9, 11, 14, 17, 19, 22	90	5
Scan time (min)	~4	2:30–4:30				~17:00	~18:00	~4:00

T1w, T1-weighted; T2w-FLAIR, T2-weighted-fluid-attenuated inversion recovery; MAGiC, magnetic resonance image compilation; FOV, field of view; TR, repetition time; TE, echo time; TI, inversion time.

https://dx.doi.org/10.21037/qims-24-721

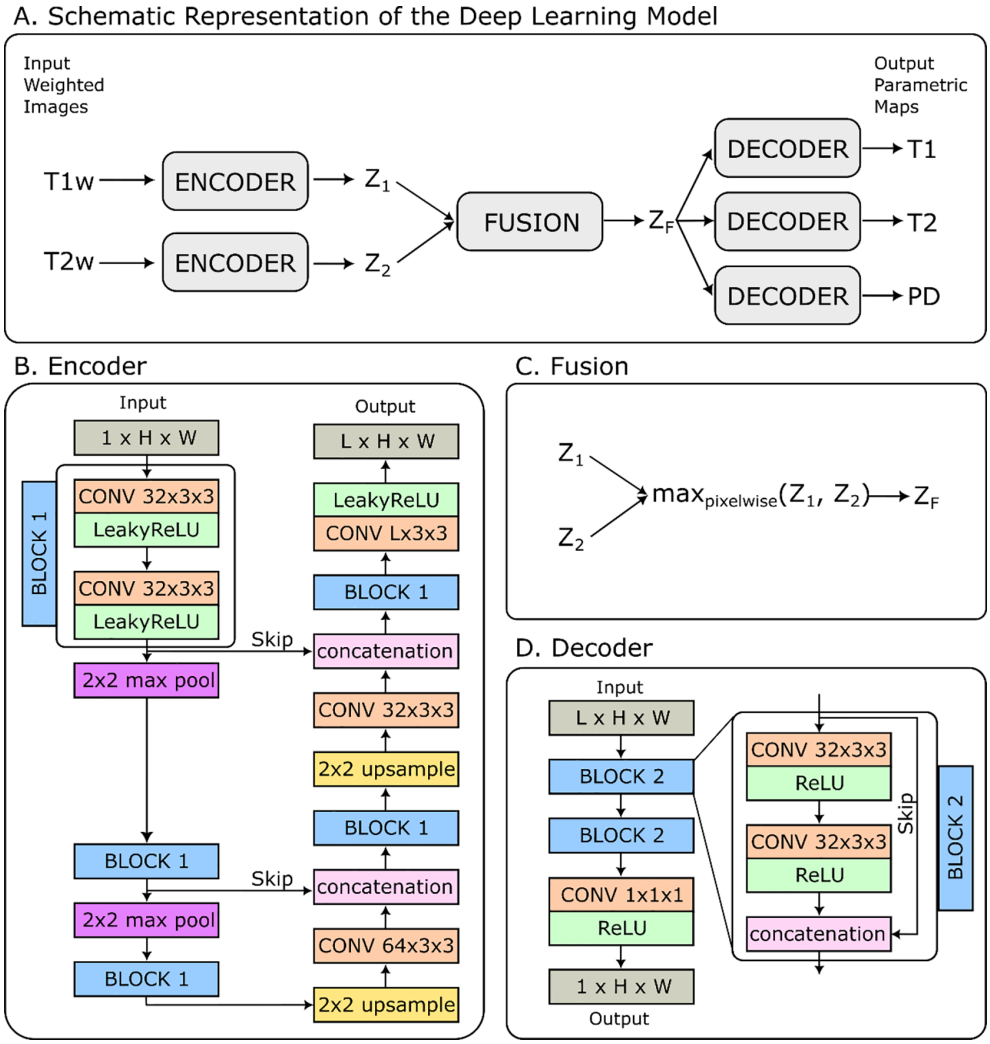

Figure S1 DL model architecture. (A) Schematic representation of the DL model. (B) Architecture of the encoder modules. (C) Fusion stage. (D) Architecture of the decoder modules. DL, deep learning.

https://dx.doi.org/10.21037/qims-24-721

Table S2 Indexes of patients selected for training, early-stopping validation and test of the GLIOMA dataset

Test	Early-stopping validation	Training
Test with GLIOMA (leave-one-out)		
0	1, 6	2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14
1	4, 11	0, 2, 3, 5, 6, 7, 8, 9, 10, 12, 13, 14
2	8, 11	0, 1, 3, 4, 5, 6, 7, 9, 10, 12, 13, 14
3	14, 11	0, 1, 2, 4, 5, 6, 7, 8, 9, 10, 12, 13
4	14, 11	0, 1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 13
5	1, 11	0, 2, 3, 4, 6, 7, 8, 9, 10, 12, 13, 14
6	14, 10	0, 1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 13
7	14, 10	0, 1, 2, 3, 4, 5, 6, 8, 9, 11, 12, 13
8	5, 6	0, 1, 2, 3, 4, 7, 9, 10, 11, 12, 13, 14
9	12, 11	0, 1, 2, 3, 4, 5, 6, 7, 8, 10, 13, 14
10	1, 6	0, 2, 3, 4, 5, 7, 8, 9, 11, 12, 13, 14
11	5, 6	0, 1, 2, 3, 4, 7, 8, 9, 10, 12, 13, 14
13	12, 6	0, 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 14
14	8, 11	0, 1, 2, 3, 4, 5, 6, 7, 9, 10, 12, 13
Test with independent datasets		
UPenn-GBM and RMaps	1, 6	0, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14

Note that patients 6, 10, and 11 do not present T1e, and patient 12 was not included in any test set because the T2w-FLAIR was not available. T1e, T1-weighted-enhancement; T2w-FLAIR, T2-weighted-fluid-attenuated inversion recovery.

https://dx.doi.org/10.21037/qims-24-721

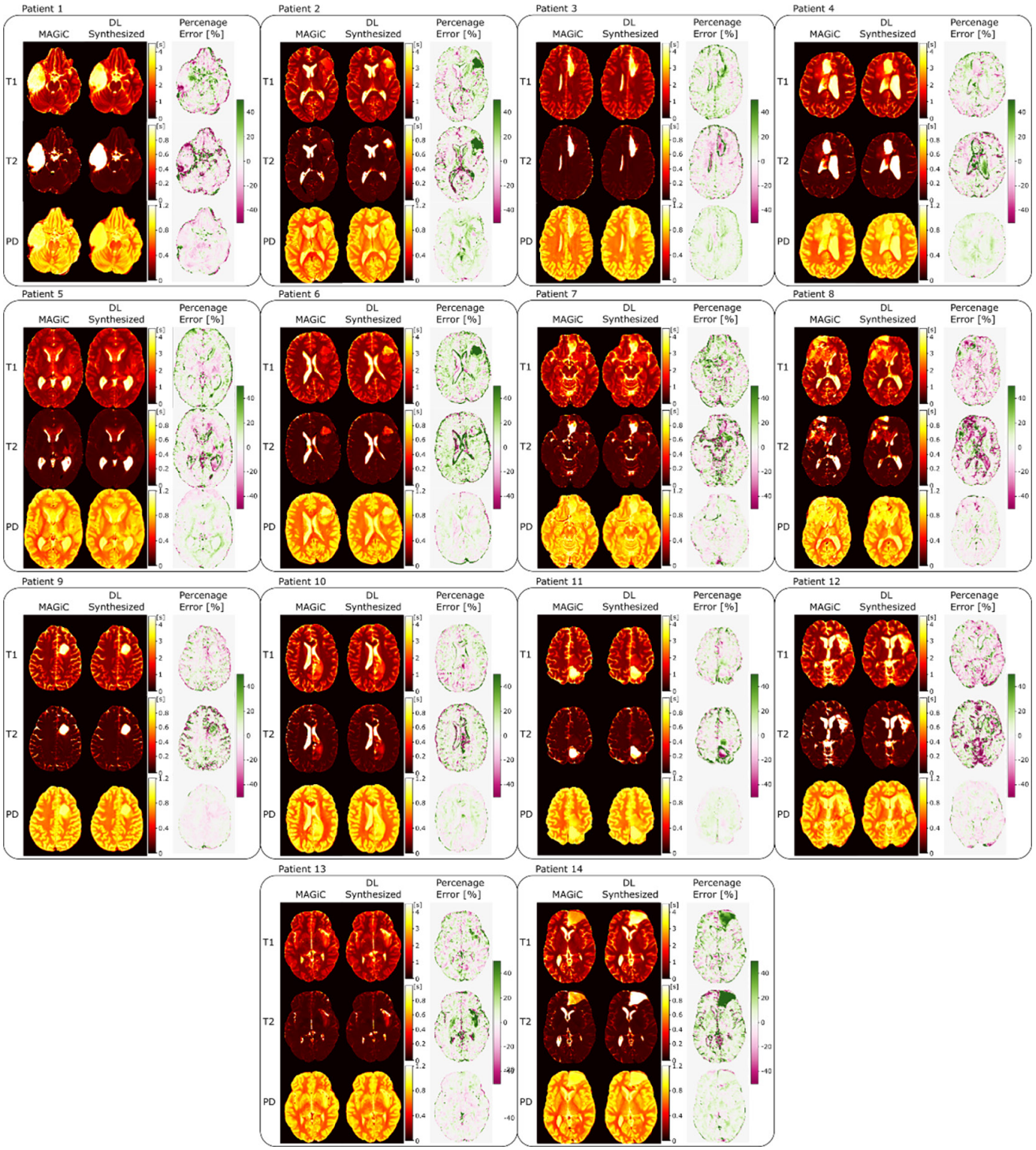

Figure S2 A representative axial slice of MAGiC and corresponding synthesized maps for the test patients of leave-one-out cross-validation with GLIOMA dataset. The voxel-wise percentage error is also represented for each pair of maps.

https://dx.doi.org/10.21037/qims-24-721

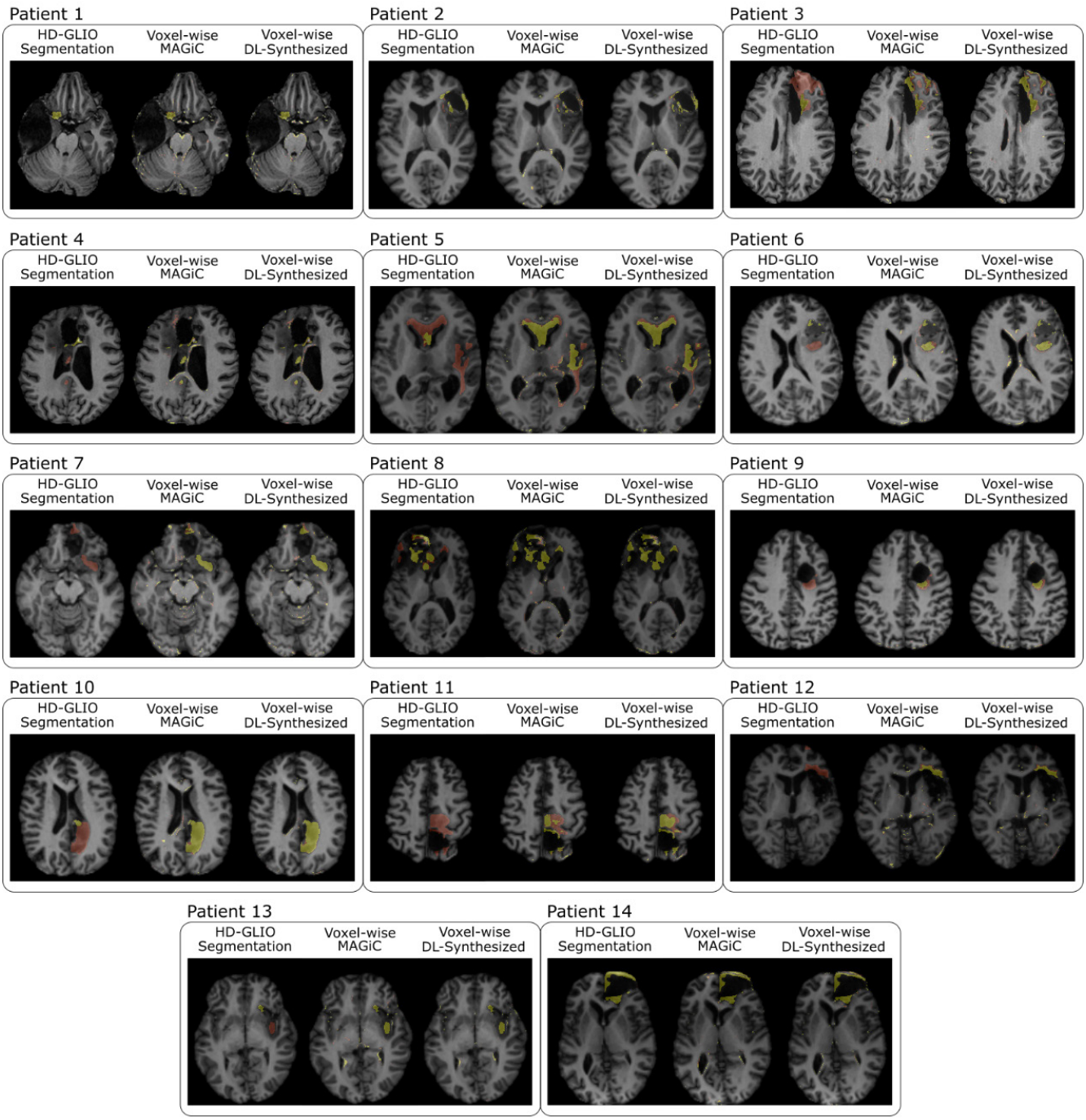

Figure S3 A representative axial slice of the T1e and T2h segmentations overlaid on the pre-T1w images for all the test patients of leave-one-out cross-validation with GLIOMA dataset. For each patient, the ground-truth segmentation obtained with HD-GLIO from the four weighted images, the segmentations obtained through the statistical predictions from MAGiC maps, and the segmentations obtained through the statistical predictions from synthesized maps are shown. MAGiC, magnetic resonance image compilation; T1e, T1w-enhancement; T2h, T2w/T2w-FLAIR signal hyperintensity; T1w, T1-weighted; FLAIR, fluid-attenuated inversion recovery.

https://dx.doi.org/10.21037/qims-24-721